OPIC
Office de la propriété intellectuelle du Canada

CIPO
Canadian Intellectual Property Office

(12) (19) (CA) **Brevet-Patent**

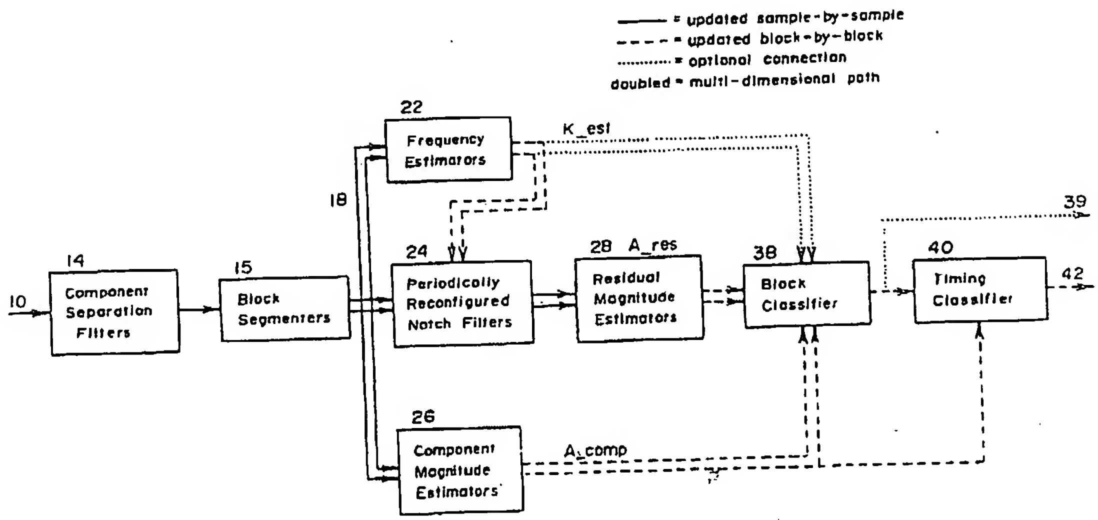
(11) (21) (C) **2,094,412**
(22) 1993/04/20
(43) 1994/06/23
(45) 1998/05/19

(72) Cox, Neil Bernard, CA
(72) Froese, Edwin Lloyd, CA
(73) Sentry Telecom Systems Inc., CA

(51) Int.Cl.[6] H04Q 1/46
(30) 1992/12/22 (07/994,910) US

(54) DETECTEUR ET CLASSIFICATION DE SIGNAUX MULTIFREQUENCE

(54) MULTI-FREQUENCY SIGNAL DETECTOR AND CLASSIFIER



(57) Méthode et appareil de détection et de classification de signaux constituant une combinaison additive de quelques composantes sinusoïdales à amplitude constante, ci-appelées tonalités N. Les attributs de cette méthode et de cet appareil permettent d'offrir une performance de classification supérieure en utilisant un algorithme de faible complexité. La méthode comprend les éléments suivants : assurer un filtrage pour enlever les composantes de signal parasites, séparer le signal d'entrée en un ou plusieurs trains de sortie puis segmenter et grouper des ensembles de blocs alignés dans le temps, sans chevauchement, d'échantillons de données successifs, ou des ensembles de blocs. On obtient ensuite les ensembles de blocs alignés dans le temps en regroupant tous les blocs ayant les mêmes temps de démarcation et en sélectionnant ainsi un bloc de chaque train. Le processus suivant est appliqué dans tous

(57) A method and apparatus for detecting and classifying signals that are the additive combination of a few constant-amplitude sinusoidal components, herein called N-tones. The attributes of this method and apparatus include provision of superior classification performance with an algorithm of low computational complexity. The method includes filtering to remove extraneous signal components, separation of the incoming signal into one or more output streams and segmenting and grouping sets of non-overlapping time-aligned blocks of successive data samples, or block-sets. Time-aligned sets of blocks are then obtained by grouping together all blocks with the same demarcation times, thereby selecting one block from each stream. The following process is applied whenever a new block-set becomes available. The magnitude of the data within each block of a block-set is estimated along with the

MULTI-FREQUENCY SIGNAL DETECTOR AND CLASSIFIER

BACKGROUND

The present invention relates to a detector and classifier for signals that are the additive combination of a few constant-amplitude sinusoidal components, hereafter called N-tones. The invention herein has particular application to the sub-class of N-tones called dual-tone multi-frequency (DTMF) telephone signals.

DTMF signals are N-tones used for representing telephone numbers and other signalling functions within the telephone system. Detailed specification of DTMF signal properties have been standardized by international agreement. Sixteen unique DTMF signals are defined; one for each of the numbers on a telephone keypad plus six for additional keys. Ignoring noise, distortion and allowable equipment variability, each DTMF signal is an additive combination of two equal-amplitude tones. The frequencies of the component tones serve to distinguish one DTMF signal from another. Specifically, each DTMF signal is comprised of two tones with frequencies taken from two mutually-exclusive frequency bands. For example, the signal generated by depressing "1" on the telephone keypad is the sum of a 697 Hz tone and a 1209 Hz tone, and the signal generated by depressing "5" is the sum of a 770 Hz tone and a 1336 Hz tone. The low frequency band, or low-band, is comprised of tones with frequencies of (nominally) 697 Hz, 770 Hz, 852 Hz and 941 Hz. The high frequency band, or high-band, is comprised of tones with frequencies of (nominally) 1209 Hz, 1336 Hz, 1477 Hz and 1633 Hz.

In telephony applications, one must be able to quickly detect and accurately classify DTMF signals that are embedded in noise, and one must not falsely indicate DTMF presence within other valid signals. The second issue generally presents the largest challenge because short segments of speech occasionally appear very DTMF-like.

Family", published in the Proceedings of the 1989 IEEE International Conference on Acoustics, Speech and Signal Processing, at pages 1134 to 1137, current standards of performance in DTMF detectors regard five false detections in a 30 minute sampling of speech as a good level of performance. The present invention produced no false detections in over 220 minutes of test material, including samples of telephone traffic, a radio talk show, music and the same 30 minute speech sampling as was used in establishing the aforementioned standard of performance. This was achieved with an algorithm that consumes a small fraction of the computing capacity of present-day digital signal processors.

BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself, however, as well as other features and advantages thereof, will be best understood by reference to the description which follows, read in conjunction with the accompanying drawings wherein:

Figure 1 is a general schematic diagram of a preferred embodiment of the invention;

Figure 2 is a schematic diagram of an embodiment of the invention configured for detection of DTMF telephony signals.

DETAILED DESCRIPTION WITH REFERENCE TO THE DRAWINGS

Referring to Figure 1 there is shown a block diagram of a representative embodiment of the invention for processing a stream of input data on input line 10. The input data is in digital form, being samples of an analog signal taken at a

set. It is assumed in the sequel that the block segmenters 15 are configured for generating block-sets as described above.

The block-sets are directed over the output streams 18 to the frequency estimators 22, the periodically-reconfigured notch filters 24 and the component magnitude estimators 26. The outputs of the notch filters 24 are directed to residual magnitude estimators 28. The purpose of these elements is to derive data required for determining if input line 10 has the requisite properties of an N-tone.

The average magnitude for each block within a block-set is estimated by the corresponding component magnitude estimator 26, thereby producing one L-dimensional vector of component magnitude estimates A_comp, where L is the number of parallel streams within inputs 18. These average magnitude estimates can be computed by simply summing the square of the sample values in each block.

The frequency estimators 22 are used to identify the frequencies of the dominant spectral tones so that the notch filters may be initialized to filter out these tones. Specifically, the frequency estimators 22 act on the block-set to estimate the N-dimensional vector of frequency-determining filter coefficients K_est, where N is the number of tonal components in the N-tone. This frequency estimation can be performed using one of a number of methods for estimating the frequency of sinusoids embedded in noise that have been described in the literature. A method based on linear prediction is a suitable choice.

particularly when $\beta \approx 1$. By superposition it follows that such ringing will occur whenever there is a change in the amplitude of such a tone. The implication of this ringing phenomenon combined with the periodic filter reconfiguration is that the output magnitude of a constrained notch filter remains low relative to its input only if the input is dominated by a constant-amplitude tone that spans the entire length of the block of data under analysis. Such a restriction is seldom met by non-stationary signals such as speech.

The requirements for the component separation filters 14, the frequency estimators 22 and the notch filters 24 are interdependent. It was stated earlier that the component separation filters produce one or more isolated streams which, in the presence of an N-tone, each contain one or more (i.e. J) tonal components. The simplest filtering option is to produce only one output stream. However, no information is then provided to the block classifier about the relative strength of each tonal component unless further component separation is performed during application of the notch filters 24. One may consider using a cascade of lower-order notch filters for multi-component streams in order to obtain additional information about individual signal components, but special measures are then needed to ensure that transient effects of the periodic filter reconfiguration do not significantly affect later filters in the cascade. One method of minimizing these effects is to delay the periodic reconfiguration for later filtering stages until the transients have largely disappeared. However, this leads to an increase in the variance of the residual magnitude estimates 28 because fewer data points are available for analysis. Finally, while a number of well-known techniques can be used for estimating component frequencies when the number of components is known, these techniques are simpler and more efficient when the number of components is small.

long block-size also takes better advantage of non-stationarity in signals such as speech, thereby reducing the likelihood of misclassification. On the other hand, short block-sizes reduce the delay in detecting an N-tone and facilitate detection of short-duration N-tones.

An embodiment for DTMF Detection and Classification

Referring to Figure 2 there is shown a block diagram of an embodiment of the invention configured for detection and classification of DTMF telehony signals. Like reference numbers as in Figure 1 are used to refer to like parts. The entire assembly was implemented on a 27 MHz DSP56001 digital signal processor. The block diagram includes a two-to-one downsampler 12 and an input 10 to which is connected a stream of sampled data that is sampled at 8000 Hz. This downsampling involves simply ignoring every other sample. Downsampling is performed to reduce the processing requirements and to reduce the required sharpness of the subsequent band-isolation filters.

The downsampled data on line 13 is directed to the high-band isolation filter 14H and the low-band isolation filter 16. The high-band isolation filter 14H isolates the high-band DTMF tones (>1200 Hz) from the low-band DTMF tones (<1000Hz). A 5th-order elliptic infinite impulse response (IIR) high-pass filter designed for a sampling frequency of 4000 Hz, band-edge frequency of 1160 Hz, 0.25 dB pass-band ripple and 40 dB stop-band attenuation is used. The low-band isolation filter 16 isolates the low-band DTMF tones from dial tone (<500 Hz) and high-band DTMF tones. An 8th-order elliptic IIR band-pass filter designed for a sampling frequency of 4000 Hz, band edges of 630 Hz and 1010 Hz, 0.25 dB pass-band ripple and 40 dB stop-band attenuation is used.

The output streams from the filtering stage are each passed through a block segmenter 15 and processed to derive the

where x[] is a block of successive data samples and BS is the number of samples in the block. For the present embodiment it is not necessary to explicitly derive the component frequencies, rather, it is sufficient to proceed only as far as is required to estimate K_est_j. It is also only necessary to produce estimates of K_est_j once every BS samples.

The aforementioned formula for K_est_j is sometimes applied twice within each frequency estimator with different values of K in order to minimize both the variance and the bias of the derived estimate. Estimation variance increases as the denominator of the expression (r[k]) decreases. From known properties of the autocorrelation function this suggests that the best choice of k is zero. Unfortunately, the result when k=0 can be biased when noise is present because of the positive contribution that noise makes to r[0]. The solution is to initially apply the K_est_j estimator with a compromise value of k that is relatively effective for all possible tones, and then re-apply the estimator with a new value of k if a better one exists for the tone that is present. For the DTMF detector illustrated in Figure 2, initial compromise values of k=5, and k=2 are suitable for data streams 18H and 18L, respectively. Based on the initial estimates of K_est_L, one then chooses offsets of 3, 5, 5 and 2 when the tone is near the nominal low-band DTMF frequencies of 697 Hz, 770 Hz, 852 Hz and 941 Hz respectively. Similarly, for K_est_H one should use offsets of 5, 3, 4 and 5 for the second pass in the presence of 1209 Hz, 1336 Hz, 1477 Hz or 1633 Hz tones, respectively.

The notch filters 24H and 24L are each periodically-reconfigured second-order constrained notch filters. The associated system function and the details for periodic reconfiguration were presented in the discussion of Figure 1. Choosing ß =0.87 provided a convenient tradeoff between speech

40 ms shall be accepted, 3) non-DTMF segments of less than 20 ms within a DTMF segment shall be ignored, and 4) non-DTMF segments of greater than 30 ms shall be recognized. The present implementation uses two timing thresholds to impose these restrictions: one for acceptance or rejection of DTMF-like segments, and one for acceptance or rejection of a non-DTMF segment. The recommended thresholds are 36 ms and 26 ms, respectively.

The timing classifier 40 is driven by the time course of block classifications 39 and by the low-band component magnitude estimate A_comp_L. The timing classifier 40 is considered to be in a stable state whenever each block classification 39 agrees with the most recently asserted output class 42. The first block classification that contradicts the asserted output class throws the program into a controlled race condition, where the estimated duration of DTMF and non-DTMF are simultaneously accumulated. The race winner is the first duration to reach its timing threshold. If the race is won by the same class as was previously asserted, then one returns to the stable state without altering output class 42. Otherwise, the indicated class change is conveyed to the output class 42 prior to returning to the stable state.

Certain properties of the block classifier's output result in the need for refined estimates of signal duration within the timing classifier 38. Except for some relatively minor exceptions, a DTMF signal is not detected by the block classifier 40 unless it completely fills the block-set under analysis. If simple block counts are used as duration estimates, then the results depend on the coincidental alignment between the block demarcation boundaries and on/off transitions in the DTMF signal. For example, if the DTMF signal is 30 ms and the analysis block-size is 15 ms, then the analysis may separate the DTMF signal into two completely-filled blocks or it may separate

The following additional conditions were included in the timing classifier 40 for proper performance. Firstly, a one block delay is built into the non-DTMF duration counter so that transitions from non-DTMF to DTMF are properly handled, that is, when the current block is classified as non-DTMF, one needs to know whether the next block is DTMF-like or non-DTMF before it is possible to determine how much of the current block is non-DTMF. Secondly, interruption of a string of non-DTMF blocks by a DTMF-like block causes reset of the non-DTMF duration counter. This ensures that short gaps caused by the repeated bounce of a switch do not accumulate and erroneously appear to be a valid inter-digit pause. Finally, a change from one DTMF class to another during a race must cause reset of the DTMF duration counter. The new DTMF class will then be asserted only after it alone is present for a sufficient amount of time.

Performance of the Embodiment for DTMF Classification

The performance of the block classifier 38 for DTMF signal detection is optimistically described by two assertions. The first assertion is the output class 39 is asserted to be DTMF-like only when the block is full of DTMF signal. Tests have shown this always to be true at the start of a DTMF signal. However, a block which straddles a DTMF signal's endpoint may be classified as DTMF-like if the block's endpoint is less than about 3 ms beyond the DTMF signal's endpoint. The second optimistic assertion is that the block classifier's output 39 is always asserted to be DTMF-like when the block is filled with a DTMF signal. Tests have shown that this may be false if the leading edge of the block is within about 4 ms of the onset of the DTMF digit. This is due to the aggregate transient response of the channel and the input filters, which "smears" the startup transition across a number of samples, thereby producing ringing in the constrained notch filters.

While this invention has been described with reference to illustrative embodiments, this description is not intended to be construed in a limiting sense. Various modifications of the illustrative embodiments, as well as other embodiments of the invention, will be apparent to persons skilled in the art upon reference to this description. It is therefore contemplated that the appended claims will cover any such modifications of embodiments as fall within the true scope of the invention.

(h) means for testing the frequencies of the identified dominant spectral components to partially ascertain if the corresponding signal on said source line has the requisite properties of an N-tone;

(i) means for deriving a running estimate of the duration of time that the signal on said source line is an N-tone;

(j) means for deriving a running estimate of the duration of time that the signal on said source line is not an N-tone;

(k) means for testing the durations of candidate N-tones and intervening non-N-tone segments for conformance with signal persistence and longevity requirements.

2. Apparatus according to claim 1, wherein said rejecting means includes a digital filter for each of said output streams.

3. Apparatus according to claim 2, wherein said rejecting means include a means of sample-rate conversion.

4. Apparatus according to claim 1, wherein said means for identifying the dominant spectral components employs linear prediction analysis to produce as output the set of frequency-determining filter coefficients defined by

$$K\_est_j = 2\cos(2\pi f\_est_j\, T)$$

where K_est_j is the estimated frequency-determining filter coefficient for component j, f_est_j is the frequency identified by K_est_j and T is the sampling period for signals on each of said output streams.

8. Apparatus according to claim 5, wherein said function employs a squaring operation.

9. Apparatus according to claim 1, wherein said means for deriving a running estimate of the duration of an N-tone employs the formula:

$$dur = BS\ (A_{C+1}/A_C + C + A_0/A_1)$$

when the same N-tone has been detected in each of the preceding C block classifications but the most recently derived block classification was not an N-tone, and

$$dur = BS\ (A_{C+1}/A_C + C + 1)$$

when the same N-tone has been detected in each of the preceding C block classifications and the most recently derived block classification was also the same N-tone, and dur = 0 when neither of the above conditions apply, where A_i is a component magnitude estimate from I blocks preceding the most recently acquired set of blocks and BS is the duration of the block.

10. Apparatus according to claim 7, wherein a change from one N-tone class to another causes said running estimate of the duration of the N-tone to reset.

11. Apparatus according to claim 1, wherein said means for deriving a running estimate of the duration of time that the signal of said source line is not an N-tone is a compilation of the time not counted as an N-tone for said running estimate of the duration of an N-tone.

components from each said block from each said output stream;

(f) a residual magnitude estimator coupled to the output of said notch filter and operative for estimating the magnitude of data within each said block from each said output stream after the identified dominant spectral components have been removed;

(g) a block classifier coupled to outputs of all component magnitude estimators, all frequency estimators and all residual magnitude estimators, and operative to test if the signal on said source line has the requisite properties of tonal purity, absolute component magnitude, relative component magnitude and component frequency, the output of said block classifier providing an indication of the class of N-tone is present during times when said requisite properties are satisfied, the output of said block classifier also providing an indication of classification failure during times when said requisite properties are not satisified;

(h) a timing classifier coupled to the output of said block classifier operative for testing the duration of a candidate N-tone for conformance with application-specific requirements regarding signal persistence and longevity, and other signal timing restrictions as required by the application; and

(i) means for coupling a subset of the output of said component magnitude estimators to said timing classifier so that refined estimates of signal duration may be computed by said timing classifier.

both the purity and the frequency of said dominant spectral components.

17. Apparatus according to claim 13, wherein said component magnitude estimator and said residual estimator each estimate the magnitude of data within each said block of data using the average of a non-linear function of data samples within said block.

18. Apparatus according to claim 17 wherein said non-linear function is a squaring operation.

19. Apparatus according to claim 13, wherein said timing classifier employs the following formula to derive a running estimate of the duration of an N-tone:

$$dur = BS\cdot(A_{C+1}/A_C + C + A_0/A_1)$$

when the same N-tone has been detected in each of the preceding C block classifications but the most recently derived block classification was not an N-tone, and

$$dur = BS\cdot(A_{C+1}/A_C + C + 1)$$

when the same N-tone has been detected in each of the preceding C block classifications and the most recently derived block classification was also the same N-tone, and dur=0 when neither of the above conditions apply, where A_i is a component magnitude estimate from I blocks preceding the most recently acquired set of blocks and BS in the number of samples in each block.

20. Apparatus according to claim 13, wherein said timing classifier includes means for deriving a running estimate of the duration of time that the signal on said source line is an N-tone and wherein a change from one N-tone class to another causes said running estimate of the duration of the N-tone to reset.

(f) estimating the magnitude of data within each said block from each said output stream after removal of the identified dominant spectral components;

(g) testing the magnitude estimates that were obtained both before and after removal of the identified dominant spectral components, thereby partially ascertaining if the signal on said source line has the requisite properties of an N-tone;

(h) testing the frequencies of the identified dominant spectral components, thereby partially ascertaining if the signal on said source line has the requisite properties of an N-tone;

(i) deriving a running estimate of the duration of time that the signal on said source line is an N-tone;

(j) deriving a running estimate of the duration of time that the signal on said source line is not an N-tone;

(k) testing the durations of candidate N-tones and intervening non-N-tone segments for conformance with signal persistence and longevity requirements.

24. A method according to claim 23, wherein said rejecting step of extraneous signal components utilizes a combination of digital filters, downsampler and sample-rate converters.

25. A method according to claim 23, wherein said said step of identifying of dominant spectral components includes employing linear prediction analysis to produce as

filters and the setting of K_est_j prior to each application of the filter to a block of data, the values of K_est_j for said periodic reconfiguration being provided by the output of said frequency estimator after being applied to the same block of data.

28. A method according to claim 23, wherein said estimating of the magnitude of data within a block of data, either before or after removal of the dominant spectral components, includes the averaging of a function of data samples within said block.

29. A method according to claim 28, wherein said function includes a squaring operation.

30. A method according to claim 23, wherein said derivation of a running estimate of the duration of an N-tone includes the formula:

$$dur = BS \cdot (A_{C+1}/A_C + C + A_0/A_1)$$

when the same N-tone has been detected in each of the preceding C block classifications but the most recently derived block classification was not the same N-tone, and

$$dur = BS \cdot (A_{C+1}/A_C + C + 1)$$

when the same N-tone has been detected in each of the preceding C block classifications and the most recently derived block classification was also the same N-tone, and dur=0 when neither of the above conditions apply, where A_i is a component magnitude estimate from i blocks preceding the most recently acquired set of blocks and BS is the number of samples in each block.

31. A method according to claim 23, including resetting said running estimate of the duration of the N-

ABSTRACT

A method and apparatus for detecting and classifying signals that are the additive combination of a few constant-amplitude sinusoidal components, herein called N-tones. The attributes of this method and apparatus include provision of superior classification performance with an algorithm of low computational complexity. The method includes filtering to remove extraneous signal components, separation of the incoming signal into one or more output streams and segmenting and grouping sets of non-overlapping time-aligned blocks of successive data samples, or block-sets. Time-aligned sets of blocks are then obtained by grouping together all blocks with the same demarcation times, thereby selecting one block from each stream. The following process is applied whenever a new block-set becomes available. The magnitude of the data within each block of a block-set is estimated along with the frequencies of the dominant spectral components. The frequency estimates are then used as part of the configuration process of a set of notch filters for removal of the identified dominant spectral components. The newly-configured notch filters are then applied to the same data blocks and the magnitudes of their outputs are estimated. The estimated magnitudes from before and after the notch filtering are then passed on to a block classifier along with the aforementioned frequency estimates. This block classifier tests its inputs to ascertain if the current block-set conforms to prespecified conditions of tonal purity, absolute component magnitude, relative component magnitude and component frequency. Finally, a timing classifier is applied to monitor the time course of block classifications and test the candidate N-tones for conformance with application-specific requirements regarding signal persistence and longevity.

BEST AVAILABLE COPY

2094412

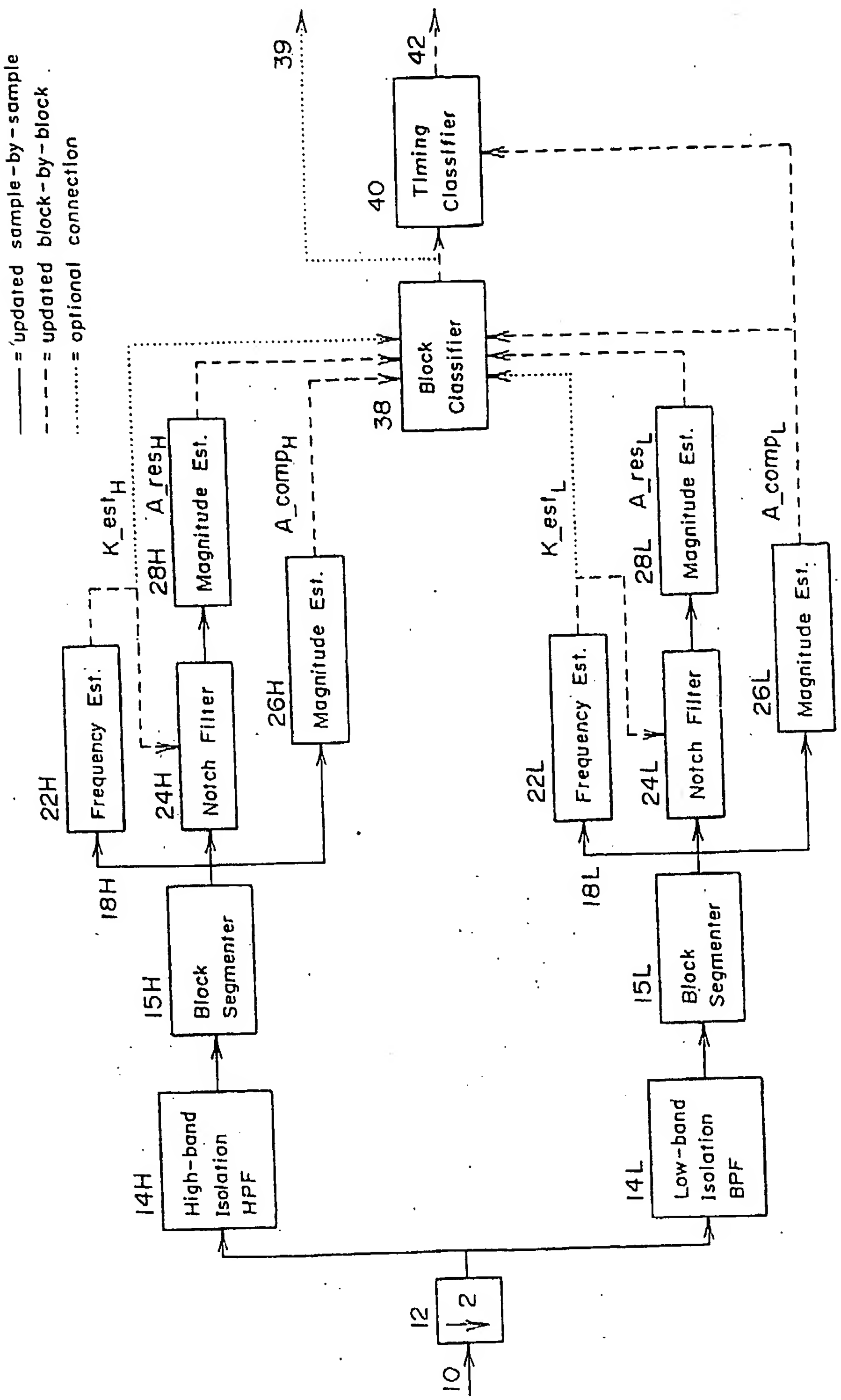


Fig. 2